

Analyzing the Vulnerability of Critical Infrastructure to Attack, and Planning Defenses

Gerald Brown
Matthew Carlyle
Javier Salmerón
Kevin Wood

Operations Research Department
Naval Postgraduate School
Monterey, California 93943

November 2005

Appears as: “Analyzing the Vulnerability of Critical Infrastructure to Attack , and Planning Defenses,” in *Tutorials in Operations Research* , INFORMS, ISBN-1-877640-21-2 pp. 102-123.

Abstract

We describe new bilevel programming models to (1) help make the country’s critical infrastructure more resilient to attacks by terrorists, (2) help government and business plan those improvements, and (3) help influence related public policy on investment incentives, regulations, etc. An intelligent attacker (terrorists) and defender (us) are key features of all these models, along with information transparency: These are Stackelberg games as opposed to two-person zero-sum games. We illustrate these models with applications to electric power grids, subways, airports, and other critical infrastructure. For instance, one model identifies locations for a given set of electronic sensors that minimize the worst-case time to detection of a chemical, biological or radiological contaminant introduced into the Washington, DC subway system. The paper concludes by reporting insights we have gained through forming “red teams,” each of which gathers open-source data on a real-world system, develops an appropriate attacker-defender or defender-attacker model, and solves the model to identify vulnerabilities in the system or to plan an optimal defense.

1 The Problem

What is critical infrastructure? The National Strategy for Homeland Security deems 13 infrastructure sectors critical to the United States; see Table 1 (DHS 2002). These include sectors such as “Government” and “Public Health,” but a number, such as “Transportation,” and “Information and Telecommunications,” comprise physical systems that connect components of our economy: In essence, they enable the transfer and distribution of our economy’s “life forces.” We focus on defending this type of infrastructure from attacks by terrorists, but we believe almost any type of critical infrastructure deserves analysis with the techniques we describe.

| | |
|-------------------------|------------------------------------|
| Agriculture | Food |
| Water | Public Health |
| Emergency Services | Government |
| Defense Industrial Base | Information and Telecommunications |
| Energy | Transportation |
| Banking and Finance | Chemical Industry |
| Postal and Shipping | |

Table 1. Thirteen infrastructure sectors critical to the United States, as defined by the Department of Homeland Security (2002).

Any critical infrastructure system represents a huge investment of our nation’s wealth. And, minor disruptions to such a system’s components—these disruptions can be random or deliberate—can severely degrade its performance as well as the performance of dependent systems. For instance, a massive power outage can result from the failure of just a few key lines and protective circuit breakers (US-Canada Power System Outage Task Force 2003). The direct effect is to interrupt the energy supply to residential and industrial customers, but all other infrastructure systems listed in Table 1 will be affected if the power outage lasts long enough. So, how do we carry out a “vulnerability analysis” when terrorist attacks are the key concern?

That is, how do we analyze the vulnerability of a critical infrastructure system to a terrorist attack, or set of coordinated attacks, and make informed proposals for reducing that vulnerability?

Most infrastructure systems are engineered to handle disruptions that result from accidents, or from random acts of nature, with little or no degradation in performance. Real-time reliability assessment of an electric power grid pronounces the system robust if no crippling “single point of failure” exists (e.g., Wood and Wollenberg 1996). Analysts of transportation systems, power plants and other infrastructure often use fault trees to assess vulnerability (Roberts et al. 1981). Such an assessment helps identify minimal sets of events, or “cutsets,” that are most likely to disrupt the system, and pronounce the system robust if their combined probability is sufficiently low. This assessment can suggest changes to the system to improve robustness, and the overall methodology can be used to evaluate alternative system configurations proposed by the analyst.

However, infrastructure that resists single points of random failure, or whose cutsets have low occurrence probabilities, may not survive a malicious, intelligent attack. For example, a lone attacker with a high-powered rifle could gravely damage an entire electric power grid by targeting highly reliable components at just a few key substations. (We reach this conclusion from our own analyses of electric power grids and from reports of gunfire disabling a substation; see Wallace 1998.) And, cutsets that are likely to occur due to random causes may not share any similarities to the cutsets that an attacker will likely find. An analyst might attempt a fault-tree assessment of a system subject to attack by guessing at the probability each individual component might be attacked. In fact, such analysis is practiced (Garcia 2001, pp. 39-48), but the results must be classified as guesses. We require a new paradigm for vulnerability analysis.

The new paradigm must account for an adversary's ability to collect information about an infrastructure system and use that information to identify weak spots in the system's architecture. A captured Al Qaeda training manual (Department of Justice 2004) advises: "Using public sources openly and without resorting to illegal means, it is possible to gather at least 80% of information about the enemy." We interpret that statement to mean: "It is possible to gather, from public sources, at least 80% of the information needed to plan a highly disruptive attack on an infrastructure system." Our experience indicates that one can often find all the information necessary to plan such an attack.

Our backgrounds compel us ask how a military analyst, faced with an intelligent enemy, would approach vulnerability analysis for military infrastructure. First, the analyst would assume that our infrastructure will be attacked and that we must take steps to protect it, i.e., harden the infrastructure or improve its active defenses. The budget for hardening or actively defending infrastructure will always be limited. So, typically, the analyst would be instructed to create a prioritized list of "defended assets" in most need of protection, along with a list of potential defensive measures, and deliver those lists to higher-level decision makers. The latter parties would make the final decisions after balancing costs, effectiveness, and intangibles, and after determining the true budget (which may be monetary or may be the number of aerial sorties, cruise missiles, tanks, etc., that can be spared for defensive purposes). Table 2 shows the doctrinal components that the US Army uses to guide the prioritization of its defended assets (as well as its enemys').

Criticality

How essential is the asset?

Vulnerability

How susceptible is the asset to surveillance or attack?

Reconstitutability

How hard will it be to recover from inflicted damage, considering time, special repair equipment, and manpower required to restore normal operation?

Threat

How probable is an attack on this asset?

Table 2. Criteria for prioritizing defended assets (Department of the Army 2002a,b).

Any person who has had a course in discrete optimization understands the fundamental flaw in the concept and use of a prioritized list. In addition to that shortcoming of the nominal military approach, we see that the civilian problem itself differs from the military one:

- Almost every civilian US asset is susceptible to surveillance or attack, and is thus vulnerable;
- No matter how hard it is to recover from inflicted damage, we will, eventually, reconstitute and recover; and
- Military planners have vast experience in determining the likelihood of alternative attacks; homeland-security planners do not. Thus, we must plan for what is possible, rather than what subjective assessments indicate is likely.

In fact, we do not try to measure the importance, or value, of an asset directly. Rather, we model a complete infrastructure system, its value to society, and how losses of the system's components reduce that value, or how improvements in the system mitigate against lost value. The exact meaning of value will depend on the system under investigation: It may mean economic output, time to detection of a toxic substance, etc., and sometimes cost, the converse of

value, will be a more convenient yardstick.

Al Qaeda teaches as its primary mission “overthrow of godless regimes (by) gathering information about the enemy, the land, the installations, and the neighbors, ... blasting and destroying the places of amusement, ... embassies, ... vital economic centers, ... bridges leading into and out of cities,” (Department of Justice 2004). Al Qaeda may not have a perfect model of a particular infrastructure system, but its operatives are instructed to gather (widely available) information about it. Clearly, that information is being used to plan the worst attacks al Qaeda can devise. Consequently, prudence dictates the assumption that al Qaeda, or any other terrorist organization, will use its limited offensive assets to maximize damage to the infrastructure system it decides to attack, and has all the data necessary to do this.

Our paradigm of an *attacker-defender model* does address criticality, vulnerability, reconstitutability, and threat, but in a very different way than military planners might. We incorporate reconstitutability by modeling how system components are repaired over time and how a repaired component contributes to improved system value (Salmeron et al. 2004b). Unless strictly defended or hardened, every system component is assumed vulnerable. We address “threat” by positing different levels of offensive resources for the terrorists. At the end of our analysis, we can determine the criticality of a group of system components, i.e., the value of protecting them, hardening them, or the value of adding new components into the system for purposes of redundancy. Another paradigm, discussed later, directly identifies an optimal defense plan: This is the *defender-attacker model*.

To understand our approach, the reader must understand the basics of the next two sections. However, a reader not interested in the mathematics may feel free to skim those details.

2 Attacker-Defender Models

The core of an attacker-defender model is an optimization model of an infrastructure system whose objective function represents the system’s value to society while it operates, or the cost to society when the system loses functionality. For instance, the maximum throughput of an oil pipeline system might measure that system’s value, while power-generation costs, plus economic losses resulting from unmet demand, might measure the full cost of operating an electric power grid.

To set the mathematical context, we assume that the defender operates a system so as to minimize cost that can be represented by a linear function. The defender’s problem is

$$\min_{\mathbf{y} \in Y} \mathbf{c} \mathbf{y}, \tag{1.1}$$

where (i) \mathbf{c} defines a vector of costs (and/or penalties), (ii) \mathbf{y} represents system operating decisions or activities, and (iii) the set Y represents constraints on that operation and the requirements to be met, e.g., road capacities in a road network, the number of commuters wishing to travel between various points in that network, etc. Of course, by including auxiliary variables in \mathbf{y} , and auxiliary constraints in Y , we can also represent certain non-linear cost functions.

We note that “defender” is actually a misnomer in these models, because the models do not directly represent defensive actions; better terms might be “system user” or “system operator.” However, our ultimate goal is to help identify defensive actions for the system user, so we feel justified in the slight abuse of terminology.

Now, our model posits that an attacker wishes to maximize the defender’s optimal (minimal) operating cost, and will do so by restricting actions \mathbf{y} . Let $x_k = 0$ if the attacker

interdicts the defender's k th asset, let $x_k = 0$ otherwise, and let \mathbf{x} denote the vector of interdiction decisions. "Interdicting an asset" may be viewed as interdicting some component of the defender's infrastructure system. For simplicity in this paper, we assume that if $x_k = 1$, then $y_i = 0$ for any activity i that requires asset k . That is, interdiction of an asset stops the defender from carrying on activities that depend on that asset. We note that defender-attacker models often exhibit a one-to-one relationship between assets and activities; for example, interdiction of a pipeline segment between cities i and j stops the single activity that can occur on that segment, "flow from i to j ."

Binary restrictions on \mathbf{x} , and some reasonable set of resource limitations on the attacker's resources, are represented by $\mathbf{x} \in X$. We represent the defender's set of feasible actions, restricted by interdictions \mathbf{x} , as $Y(\mathbf{x})$. Thus, the attacker solves this problem to guide his attacks:

$$(\text{MAX-MIN}) \quad \max_{\mathbf{x} \in X} \min_{\mathbf{y} \in Y(\mathbf{x})} \mathbf{c}\mathbf{y} . \quad (1.2)$$

MAX-MIN is a type of bilevel program (e.g., Moore and Bard 1990), which is an instance of a Stackelberg game (von Stackelberg 1952): The attacker leads with an attack and the defender follows with a response, hence the standard phrases *leader* and *follower*, for attacker and defender, respectively. The key assumption here is that the attacker has a perfect model of how the defender will optimally operate his system, and the attacker will manipulate that system to his best advantage. That is a strong but prudent assumption for the defender: He can suffer no worse if the attacker plans his attacks using a less-than-perfect model of the defender's system. We find no difficulties in assuming that the defender will operate his system optimally, but a simple adjustment to the objective function can account for certain types of inefficiencies. (More general models of inefficiency seem unsupportable. For instance, one might be able to model a defender who always operates his system at a random point along, say

the “90%-efficiency frontier,” but such a model would be hard to solve, and more importantly, to justify.)

One can devise many supportable generalizations of MAX-MIN including attacks that increase costs rather than limiting activities, or attacks that reduce the capacity of an asset less than 100%. We will cover some of these generalizations after establishing basic results.

Naturally, the defender may also lack perfect knowledge of the attacker’s capabilities. That is, the defender may be guessing at the interdiction-resource constraints contained within $\mathbf{x} \in X$. However, the defender can solve the model over a range of posited interdiction resources, and use those results to guide system improvements.

2.1 Solving an Attacker-Defender Model

For many situations, a linear program (LP) will provide an adequate model of the defender’s system and its operations. For instance, the electric power industry commonly employs linearized optimal power-flow models for security analysis (Wood and Wollenberg, p. 419). Therefore, we express the optimal operation of the defender’s system as

$$(D0) \quad \min_{\mathbf{y}} \mathbf{c}\mathbf{y} \tag{1.3}$$

$$\text{s.t. } \mathbf{A}\mathbf{y} = \mathbf{b} \tag{1.4}$$

$$\mathbf{F}\mathbf{y} \leq \mathbf{u} \tag{1.5}$$

$$\mathbf{y} \geq 0. \tag{1.6}$$

Constraints (1.4) correspond to general system-operations constraints (e.g., balance of flow in a transportation network), and constraints (1.5) represent capacity limitations for asset $i \in I$ (e.g., maximum flow across the i th network link, per unit of time). Assets might include power lines, pipelines, roads, ports, communications hubs, etc.

The attacker's interdictions might affect the system in any number of ways, but let us assume that only “assets” are in danger of being interdicted, and that interdiction of asset k causes the loss of all its capacity u_k . Thus, the full attacker-defender model is

$$\begin{aligned}
 (\text{AD0}) \quad & \max_{\mathbf{x} \in X} \min_{\mathbf{y}} \mathbf{c}\mathbf{y} \\
 & \text{s.t. } A\mathbf{y} = \mathbf{b} \\
 & \quad F\mathbf{y} \leq U(\mathbf{1} - \mathbf{x}) \\
 & \quad \mathbf{y} \geq \mathbf{0}
 \end{aligned}$$

where $U = \text{diag}(u)$. We assume that the inner LP has been constructed to be feasible for any \mathbf{x} , since we expect the system to operate in some degraded fashion for any conceivable attack. This may require the use of auxiliary variables that are not susceptible to interdiction.

A natural approach to reformulating this problem fixes \mathbf{x} temporarily, takes the dual of the inner linear program, and then releases \mathbf{x} . Unfortunately, an unappealing, nonlinear, mixed-integer program results. That model can be linearized in some instances (e.g., Wood 1993, Salmeron et al. 2004a), but an alternative model comes to mind: Change the paradigm of capacity interdiction to “cost interdiction,” and then take the dual of the inner problem. (See Cormican et al. 1998 for the mathematical details.) Specifically, let $-\mathbf{p}$ strictly bound the set of dual variables associated with $F\mathbf{y} \leq U(\mathbf{1} - \mathbf{x})$, taken over all possible values of $\mathbf{x} \in X$. Thus, p_k bounds the value of a unit of asset k 's capacity for the defender. Since we assume AD0 is feasible even when asset k has been interdicted and has no capacity, it must be possible to set a cost on asset k 's capacity that makes it too costly to use: p_k is just that cost. This is the standard approach to formulating an “elastic model”; see Brown et al. 1997 for more discussion.

Thus, AD0 is equivalent to

$$\begin{aligned}
(\text{AD1}) \quad & \max_{\mathbf{x} \in X} \min_{\mathbf{y}} (\mathbf{c} + \mathbf{x}^T P F) \mathbf{y} && [\text{Dual vars. for fixed } \mathbf{x}] \\
& \text{s.t. } A \mathbf{y} = \mathbf{b} && [\mathbf{0}] \\
& \quad F \mathbf{y} \leq \mathbf{u} && [\boldsymbol{\beta}] \\
& \quad \mathbf{y} \geq \mathbf{0},
\end{aligned}$$

where $P = \text{diag}(\mathbf{p})$. (Actually, non-strict bounds \mathbf{p} are also valid for identifying an optimal \mathbf{x} ; see Cormican et al. 1998.)

Now, when we take the dual of the inner minimization, a mixed-integer linear program (MILP) results:

$$\begin{aligned}
(\text{AD1-MILP}) \quad & \max_{\mathbf{x}, \boldsymbol{\theta}, \boldsymbol{\beta}} \mathbf{b}^T \boldsymbol{\theta} + \mathbf{u} \boldsymbol{\beta} \\
& \text{s.t. } A^T \boldsymbol{\theta} + F^T \boldsymbol{\beta} - F^T P \mathbf{x} \leq \mathbf{c} \\
& \quad \mathbf{x} \in X \\
& \quad \boldsymbol{\beta} \leq \mathbf{0}
\end{aligned}$$

See Errata, Note 1

We can solve this model directly, or using Benders' decomposition (Benders 1962). In fact, the standard Benders approach for integer \mathbf{x} begins by taking the dual of AD1-MILP with \mathbf{x} fixed, which obviously yields AD1. Thus, the max-min formulation of AD1 is a natural representation of the interdiction problem for application of Benders' decomposition.

To illustrate with a concrete, albeit simplified, example, consider the following model of a crude-oil pipeline system:

Data:

- A node-arc incidence matrix for the pipeline system
- \mathbf{b} vector of supplies and demands: $b_i > 0$ defines a supply of b_i million barrels per day (mmbbl/day) at node i , $b_i < 0$ defines a demand of b_i mmbbl/day at i , and $b_i = 0$ implies i is a transshipment node (pumping station)

- \mathbf{c}_1 vector of shipping costs by pipeline segment, i.e., arc (\$/mmbbl/day)
- \mathbf{c}_2 vector of penalties for not taking available supply (“take-or-pay penalties”) (\$/mmbbl/day)
- \mathbf{c}_3 vector of penalties for unmet demand (e.g., spot-market cost) (\$/mmbbl/day)
- \hat{I}_2 incomplete diagonal matrix with a 1 for each supply node, but 0 elsewhere
- \hat{I}_3 incomplete diagonal matrix with a 1 for each demand node, but 0 elsewhere

Variables:

- \mathbf{y}_1 flows on pipelines (mmbbl/day)
- \mathbf{y}_2 unused supply (mmbbl/day)
- \mathbf{y}_3 unmet demand (mmbbl/day)

Formulation:

$$(D_p0) \quad \min_{\mathbf{y}} \mathbf{c}_1 \mathbf{y}_1 + \mathbf{c}_2 \mathbf{y}_2 + \mathbf{c}_3 \mathbf{y}_3$$

$$\text{s.t. } A\mathbf{y}_1 - \hat{I}_2 \mathbf{y}_2 + \hat{I}_3 \mathbf{y}_3 = \mathbf{b} \quad (1.7)$$

$$I\mathbf{y}_1 \leq \mathbf{u} \quad (1.8)$$

all variables ≥ 0 .

Constraints (1.7) are elastic flow-balance constraints, and constraints (1.8) represent pipeline capacities. For simplicity, we

1. Have ignored the oil’s purchase price,
2. Assume $\mathbf{c}_2 = \mathbf{0}$ and $\mathbf{c}_1 > \mathbf{0}$,

3. Set all unmet demand penalties equal, i.e., $\mathbf{c}_3 = (c_3 \ c_3 \ \dots \ c_3)$, and
4. Assume that only pipeline segments can be interdicted (not, say, pumping stations).

Now, we proceed directly to create a cost-interdiction model in the form of AD1. Let $x_k = 1$ if the attacker interdicts asset k , let $x_i = 0$ otherwise, and let $\mathbf{x} \in X$ denote the binary restrictions on \mathbf{x} along with some plausible resource constraints. For example, intelligence indicates that the attacker can form at most T squads to carry out simultaneous attacks, so $X = \left\{ x_i \in \{0,1\} \forall i \in I \mid \sum_{i \in I} x_i \leq T \right\}$. We further note that $p = c_2$ exceeds the penalty incurred by not supplying one mmbbl/day (because $\mathbf{c}_1 > \mathbf{0}$). Thus, letting $\mathbf{p} = (p \ p \cdots p)$ and $P = \text{diag}(\mathbf{p})$, the max-min interdiction model is

$$\begin{aligned}
 (\text{AD}_p1) \quad & \max_{\mathbf{x} \in X} \min_{\mathbf{y}} (\mathbf{c}_1 + \mathbf{x}^T P) \mathbf{y}_1 + \mathbf{c}_2 \mathbf{y}_2 + \mathbf{c}_3 \mathbf{y}_3 \\
 & \text{s.t. } A \mathbf{y}_1 - \hat{I}_2 \mathbf{y}_2 + \hat{I}_3 \mathbf{y}_3 = \mathbf{b} \\
 & \quad I \mathbf{y}_1 \leq \mathbf{u} \\
 & \quad \text{all variables} \geq 0.
 \end{aligned}$$

We leave it to the reader to take the dual of the inner minimization to create ADP1-MILP, but a caveat is in order: The quality of the LP relaxation of that MILP will depend directly on how small the penalties p_i are, and the modeler may need to expend some effort in identifying small, valid values. For instance, each p_i in ADP1 can be validly reduced to $p_i - c_{1,\min} + \varepsilon$ where $c_{1,\min}$ is the smallest shipping cost a demand might incur while being satisfied, and where ε is some small, positive value.

In some instances, a cost-interdiction model like AD1 can actually be a more natural paradigm than AD0. In such cases, the analyst can avoid the AD0-to-AD1 transition and will not

have to worry about bounds on dual variables. For instance, suppose D0, with constraints (1.5) eliminated, corresponds to a shortest-path problem on a road network. In some situations, we may replace the capacity constraints by modeling the interdiction of a link k in the network as a delay d_k on the nominal length c_k . Thus, this model becomes:

$$\begin{aligned} (\text{AD}_R 1) \quad & \max_{\mathbf{x} \in X} \min_{\mathbf{y}} (\mathbf{c} + \mathbf{x}^T D) \mathbf{y} \\ & \text{s.t. } A\mathbf{y} = \mathbf{b} \\ & \mathbf{y} \geq \mathbf{0}, \end{aligned}$$

where $D = \text{diag}(\mathbf{d})$ and \mathbf{d} is the vector of delays d_k . See Israeli and Wood (2002) for details on this model and solution techniques for it. We note that $\text{AD}_R 1$ also fits into the framework of defender-attacker models, described next; section 4.4 provides an example.

3 Defender-Attacker Models

By solving an attacker-defender model, we identify a set of most critical components for an infrastructure system. This leads to some obvious heuristics for solving an “optimal defense problem,” i.e., identifying the best possible defense plan given a limited defense budget. We prefer truly optimal solutions, however.

In theory, one merely embeds the bilevel attacker-defender model in a trilevel defender-attacker-defender model (DAD) such as

$$\min_{\mathbf{z} \in Z} \max_{\mathbf{x} \in X(\mathbf{z})} \min_{\mathbf{y} \in Y(\mathbf{x})} \mathbf{c}\mathbf{y}. \tag{1.9}$$

Here, \mathbf{z} denotes a binary vector of defense decisions ($z_k = 1$ if asset k is hardened and made invulnerable, say, and $z_k = 0$, otherwise); $\mathbf{z} \in Z$ denotes the binary restrictions on \mathbf{z} together with budgetary (and possibly other) constraints; and the inner max-min problem simply represents an

attacker-defender model with a restricted set of attack strategies, $X(\mathbf{z})$. So, the goal of the defender is to identify a defense plan \mathbf{z}^* so that when the attacker solves

$$\max_{\mathbf{x} \in X(\mathbf{z}^*)} \min_{\mathbf{y} \in Y(\mathbf{x})} \mathbf{c}\mathbf{y}, \quad (1.10)$$

the benefit the attacker sees, i.e., the damage the attacker can guarantee to inflict, is minimized.

Unfortunately, these trilevel problems solve only with extreme difficulty, and no conversion to an MILP appears possible, in general. (See Israeli and Wood 2002 for more details, and for the description of one special-case solution technology.)

Fortunately, certain optimal-defense problems lend themselves to easier bilevel, defender-attacker models. The defender becomes the leader in this new Stackelberg game, so we essentially reverse the meanings of \mathbf{x} and \mathbf{y} , and make the following definitions:

Indices:

k asset the defender might want to defend, and the attacker might want to attack (this simple defender-attacker model assumes a one-to-one relationship between potentially attacked and potentially defended assets)

Data:

c_k value to the attacker of attacking undefended asset k (vector form \mathbf{c})

p_k reduction in value of attacking the defender's k th asset if that asset is defended, i.e., the attacker receives benefit $c_k + p_k$, $p_k \leq 0$ by attacking defended asset k (vector form \mathbf{p})

Variables:

$x_k = \begin{cases} 1 & \text{if the defender defends his } k\text{th asset} \\ 0 & \text{otherwise} \end{cases}$

$$y_k = \begin{cases} 1 & \text{if the attacker attacks the defender's } k\text{th asset} \\ 0 & \text{otherwise} \end{cases}$$

\mathbf{x}, \mathbf{y} vector forms of y_k and x_k , respectively

Constraints:

$\mathbf{x} \in X$ resource constraints and binary restrictions on the defender's defense plan, e.g.

$$X = \{\mathbf{x} \in \{0,1\}^n \mid G\mathbf{x} \leq \mathbf{f}\}$$

$\mathbf{y} \in Y$ resource constraints and binary restrictions on the attacker's attack plan, e.g.

$$Y = \{\mathbf{y} \in \{0,1\}^n \mid A\mathbf{y} = \mathbf{b}\}$$

Formulation:

$$\begin{aligned} \text{(DA1)} \quad & \min_{\mathbf{x} \in X} \max_{\mathbf{y}} (\mathbf{c} + \mathbf{x}^T P) \mathbf{y} \\ & \text{s.t. } \mathbf{y} \in Y. \end{aligned}$$

A simplified example illustrates. Suppose intelligence reports indicate that a terrorist organization, “the attacker,” intends to send out b teams to attack b different subway stations in a city encompassing $M > b$ total stations. Municipal authorities, “the defender,” have m teams, $m < M$, with which to defend stations; a defended station becomes invulnerable to attack. The value to the defender of station k is $c_k > 0$, and we assume the attacker assigns the same values. (If not, the defender's optimal defense plan may perform better than predicted.) Let $p_k = -c_k$; thus, if station k is defended, the attacker will gain no benefit by attacking it. This “subway-defense problem” may be formulated as

$$(\text{DA}_{\text{SUB}}1) \quad \min_{\mathbf{x} \in X} \max_y \sum_{k=1}^M (c_k + x_k p_k) y_k \quad (1.11)$$

$$\text{s.t.} \quad \sum_{k=1}^M y_k = b \quad (1.12)$$

$$y_k \in \{0,1\} \forall k \quad (1.13)$$

$$\text{where } X = \left\{ \mathbf{x} \in \{0,1\}^M \mid \sum_{k=1}^M x_k = m \right\}.$$

In general, the model DA1 and instances like $\text{DA}_{\text{SUB}}1$, are difficult to solve because the inner minimization is not an LP. Thus, no general transformation exists to convert DA1 into an MILP as we converted AD1 into AD1-MILP. This situation can be resolved several ways:

1. We decide that continuous attack effort represents a reasonable approximation of reality and convert Y to $Y_{\text{CONT}} = \{\mathbf{y} \in R_+^n \mid A\mathbf{y} = \mathbf{b}\}$ (Golden 1978).
2. The LP relaxation of Y , $Y_{\text{LP}} = \{\mathbf{y} \in R_+^n \mid A\mathbf{y} = \mathbf{b}\}$, yields intrinsically binary solutions, so a conversion from DA1 into “DA1-MILP” is, in fact, possible. This is the situation with $\text{DA}_{\text{SUB}}1$, and we invite the reader to work out the details. See Brown, Carlyle, Diehl et al. (2004) for an example involving theater ballistic missile defense.
3. Or, neither of the cases above pertains, and we really must include restriction $\mathbf{y} \in \{0,1\}^n$ in the definition of Y .

Case 3 requires special techniques to solve, but solution methods better than brute-force enumeration do exist (e.g., Israeli and Wood 2002, Brown et al. 2005). This paper focuses on the simpler case, case 2.

4 What We Have Done

A terrorist organization can learn just about everything it needs to know to plan a perfect attack on our critical infrastructure. This key insight leads us to apply attacker-defender and defender-attacker models to problems of protecting this infrastructure. This section describes a number of these models (some embedded in complete decision-support systems) along with applications.

These models reflect our experience as military planners who have been asked to help target enemy infrastructure and defend our own infrastructure such as road, communication, electric-power and pipeline networks. Most of the models have been derived in the course of our research and/or our students'. We have been fortunate to be able to test many of these models by (i) defining a hypothetical but real-world scenario, (ii) assembling a "red team" of well-trained, military officer-students to gather scenario data from strictly public sources, (iii) guiding the team in building, instantiating, and running an appropriate model, and (iv) analyzing the results.

The results are always interesting, and usually lead to valuable insights. We find cases in which a given set of attackers can do more damage than we would have predicted, or less; and sometimes the attacks do not target the "obvious" components revealed in single-point-of-failure analyses. An anecdote illustrates this last point. Suppose that a terrorist organization wants to attack and close down the operations of a specific airline, at a single airport, for the purpose of disrupting the airline's finances. Based on passenger-revenue data obtained from the Internet, a red-team analysis indicates that "city A" is the most damaging airport to strike for one, large US airline. If the terrorists can afford two strikes, cities B and C would be best (Brown, Hakola, Raffeto and Yanik 2004).

4.1 Electric Power Grids: An Attacker-Defender Model

We have produced a decision-support system called VEGA (Vulnerability of Electric Grids Analyzer) (Salmerón, Wood and Baldick 2005), which identifies an optimal or near-optimal attack (i.e., a set of coordinated attacks) on an electric power grid. VEGA also animates the system operator's optimal response to that attack. Given a scenario extracted from an electric-grid database and an assessment of the level of effort needed for an attacker to target each component, VEGA determines, and illustrates graphically, which equipment-loss patterns lead to maximal damage measured in terms of load (demand for power) that must be shed (dropped). Figure 1 depicts one of VEGA's many interface screens. We note that VEGA has been built with the intention of analyzing regional, bulk-power transmission systems as opposed to local distribution systems, but it could certainly be used in the latter case.

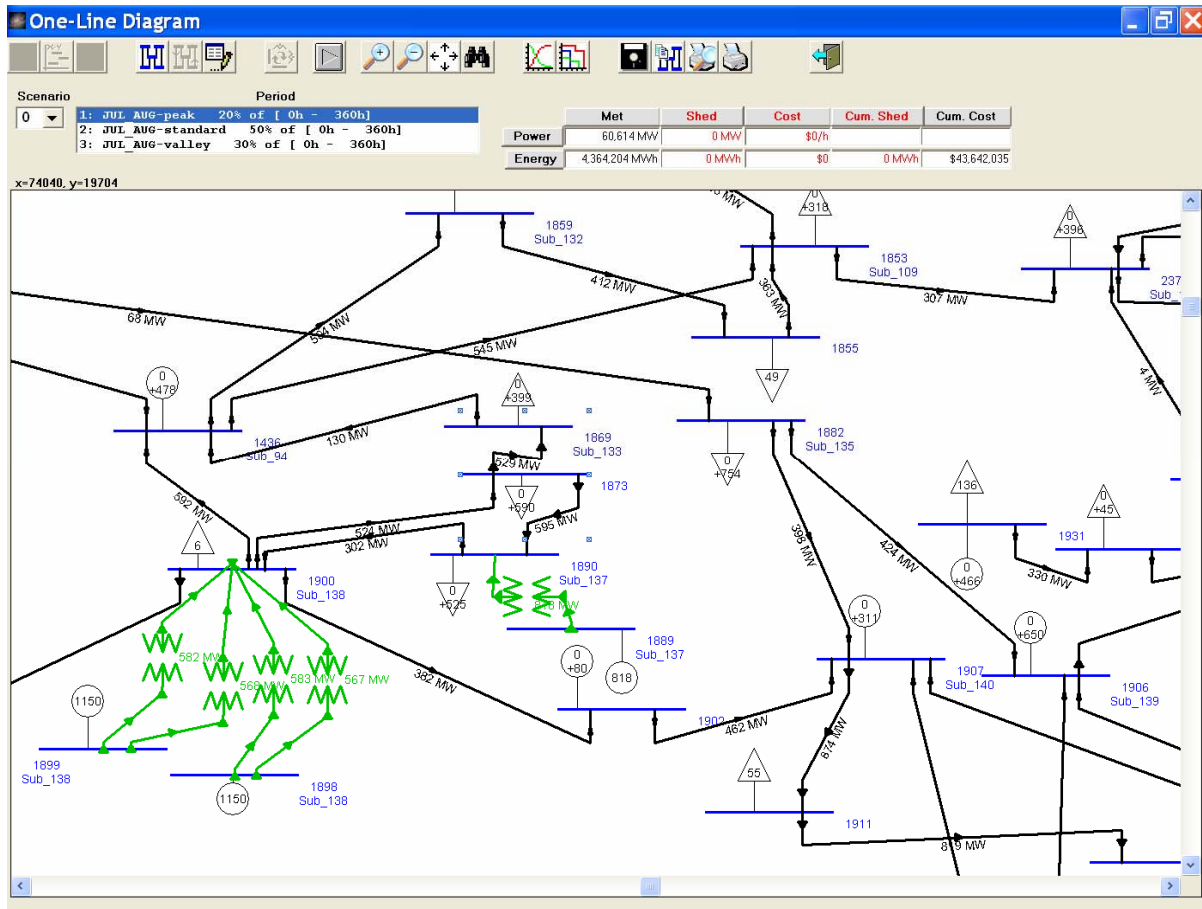


Figure 1: A screen shot of the “one-line diagram” view in VEGA (Vulnerability of Electric Grids Analyzer). The icons depict generators, transformers, buses, transmission lines, and customer demands. For any given level of attacker capability, VEGA finds an optimal or near-optimal target set of vulnerable components (Salmerón, Wood, and Baldick 2004b).

In VEGA, an “optimal DC power-flow model” comprises D0, the inner, minimizing LP. This model incorporates elastic current-balance (flow-balance) constraints along with linearized admittance constraints for AC lines. This power-flow model approximates the “true” active power flows and disregards reactive power flows, but the electric-power industry normally deems this approximation adequate for analyzing system security.

When an electric grid possesses sufficient generating and transmission capacity to meet all demand, the power-flow model reflects how a system operator would set generating levels to

minimize cost. When capacity is insufficient, as after an attack, the model reflects how the operator will react to minimize the amount of load shed, while using generation cost as a secondary criterion.

Given a fixed attack plan, VEGA must solve a sequence of power-flow models. This is true because we normally model long-term unmet demand for energy (amount of load shed, integrated over time), taking into account (i) differing repair times for components, as well as (ii) daily demand variations (“load duration curves”) and (iii) seasonal demand variations. Modeling restoration is crucial since damaged transmission lines might be repaired in a few days, other components might be repaired in a week or two, but a damaged transformer might take many months to replace. Transformers pose special difficulties because they are big, heavy and expensive; few spares exist; and a replacement might have to be ordered from, built by, and shipped from an overseas manufacturer.

An attacker-defender model can be embedded in a formal trilevel model to optimize the upgrading or hardening of a system against terrorist attack (Israeli and Wood 2002; see also Section 3 in this chapter). Such models exist for electric grids, but real-world instances are impossible to solve at this time (Salmeron et al. 2004a). Consequently, we use heuristic procedures as illustrated here. We consider a small section of the US grid containing roughly 5,000 buses, 500 generators, 3,000 loads, 5,000 lines, 1,000 transformers, 500 substations, a total reference load of 60 gigawatts (GW) and a total generating capacity of 70 GW.

We posit a group of 10 terrorists: A single terrorist can destroy a line, which takes 48 hours to repair; two terrorists can destroy a transformer or a bus which take 168 hours to repair; and three can destroy a substation, which takes 360 hours to repair. (These repair times are likely to be optimistic and serve for purposes of illustration, only.) 360 hours also represents the

study's time horizon since the system can be fully repaired in that time.

We employ a load-duration curve (a staircase function) that states: The actual load is 100% of the reference load 20% of the time ("peak load"), 70% of the reference 50% of the time ("standard load") and 45% of the reference 30% of the time ("valley load"). This load-duration curve implies a total demand for energy, over the course of the study, of about 15,000 gigawatt-hours GWh. For simplicity, we set all generation costs to \$10 per megawatt-hour (MWh) and set the cost of any unmet demand at \$1,000 per MWh.

VEGA identifies a near-optimal interdiction plan for the terrorists in about 30 minutes on a 3 (gigahertz) GHz personal computer. The plan interdicts three substations and one line, which results in 356 GWh of energy being shed over the study period, and a peak unmet load of 2.8 GW. These values are small, percentage-wise, but 2.8 GW represents the requirements of nearly three million residential customers. The economic effects of this attack would be substantial.

From these results, it is clear that protecting substations must be a priority. Therefore, we assume utility companies will spend enough money on increased security at the three hypothetically attacked substations to make them invulnerable to such attacks. We rerun VEGA with this information and find that total unmet demand reduces to less than 160 GWh and peak unmet load decreases to 1.4 GW. Once again, the terrorists attack three substations and one line.

We have reduced the disruption that the 10 terrorists can cause by about 50%, but suppose the defense budget enables us to harden the three substations attacked in the second round, plus one more: We choose one that seems to be important in a model variant that allows 15 terrorists. In the ensuing third round of attacks, the 10 terrorists attack three substations and one line, but this attack results in total unmet demand for energy of only 90 GWh and a peak unmet load of less than 600 GW. Thus, we can substantially reduce the vulnerability of this

power grid by improving security at only seven substations, from a total of roughly 500. This may be deemed cost effective by utility planners.

VEGA has been funded, in part, by the US Department of Homeland Security, Office of Domestic Preparedness, and by the Department of Energy. It uses an Intel-based computer, a Microsoft operating system and modeling software, all of which costs about five thousand dollars per seat.

4.2 Oil Pipelines: An Attacker-Defender Model

Pipeline systems for crude oil and refined petroleum products (and natural gas) are sparsely connected because of the enormous expense required to acquire right-of-ways, lay pipe, build pumping stations and maintain the system once it is complete. For instance, consider Figure 2, which is a schematic of the crude-oil pipeline network in Saudi Arabia (found, with all capacity data, through a simple search of the Internet). This network is clearly sparse, although our experience indicates that it is more densely connected than the typical gas or oil pipeline in the United States (e.g., Avery et al. 1992). In fact, the Saudi network may have substantial redundant capacity (Bremmer 2004) and, consequently, may be more resilient to attack than pipeline networks elsewhere.

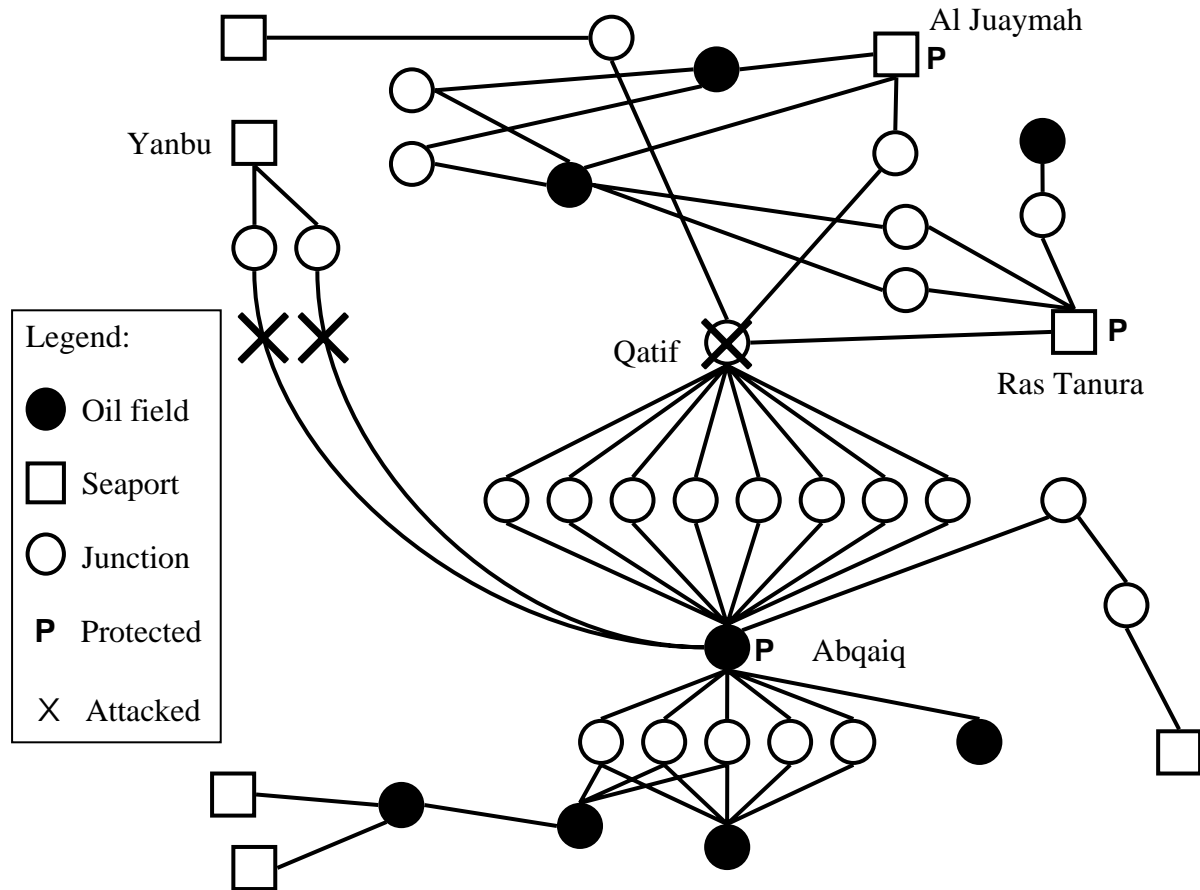


Figure 2: Three attacks on the Saudi Arabian crude-oil pipeline system reduce capacity by 3.7 mmbbl/day. The Saudi Arabian oil pipeline network has some heavily protected, invulnerable components, indicated by “P,” but most of the network is hard to defend and vulnerable to attack. Assuming insurgents have only enough resources to attack three different facilities, the three attacks shown maximally reduce Saudi capacity, even after the pipeline operator optimally redirects flows to use reserve capacity. The reduced output here exceeds a break-point estimated to cause a worldwide economic recession (Andrews et al. 2003).

An enormous security force guards the Saudi pipeline network (Sparshott 2004), but the network covers a huge area that cannot be patrolled completely. Where should the Saudi government concentrate its security efforts? A full answer to that question would require a more detailed study than we have been able to undertake. However, one of our red teams has found an “Achilles cutset,” so to speak, whose components ought to be considered closely for potential

protection.

For purposes of analysis, we play the part of a terrorist organization. First, what is our goal? Well, analysts at Morgan Stanley (Chaney and Berner 2004) report that a reduction in Saudi crude-oil output to 4 mmbbl/day (million barrels per day), from a current 8 to 9 mmbbl/day, would cause worldwide economic distress. The loss would only amount to about 5% of world demand, but Chaney and Berner estimate the price of oil would jump to \$80/bbl from a 2004 price of \$40/bbl. Furthermore, this jump could lead to a global recession if damaged facilities could not be repaired in a few months. So, taking the lead from Morgan Stanley, we set a goal of reducing Saudi oil output to 4 mmbbl/day or less.

Naturally, we would like to implement a coordinated strategy that requires as few individual attacks as possible. What is the minimum number necessary to reach our goal? We assume that the largest oil field at Abqaiq is well protected, i.e., invulnerable to attack, as are the two seaports on the Persian Gulf, Ras Tanamura and Al Juaymah. But, all other system components, pipeline segments and junctions, are potential targets.

We can solve this problem via the max-flow interdiction model of Wood (1993), which minimizes maximum flow given a fixed amount of resource. (Thus, we must solve a min-max attacker-defender model rather than paradigmatic max-min model, AD0.) Each seaport in Figure 2 is connected to a supersink, with the arc's capacity equaling the port's capacity. Similarly, each oil field is connected by an arc to a supersource, with the arc's capacity equaling the production capacity of the field. Pipeline arcs are assigned their known capacities, and junctions are split into arcs, as required, to represent limited pumping capacity.

The best single attack targets the junction at Qatif. Worldwide oil prices spike on the news, but moderate quickly when it is learned that maximum output has only been reduced to 8.7

mmbbl/day, in a system whose current total capacity is about 10 mmbbl/day, with current output around 9 mmbbl/day. (Exact values for these numbers would depend on when the hypothetical attack occurs. These values are close to current numbers, but Saudi Arabia may add capacity in the near future, and demand could increase or soften.)

The best attack on two targets adds one of the pipelines connecting Abqaiq and Yanbu, and reduces maximum output to 5.8 mmbbl/day. The world gets really worried. The best attack on three targets adds the second Abqaiq-Yanbu pipeline, Saudi output drops to 3.7 mmbbl/day, our goal has been reached, and world-wide oil prices shoot skyward.

This situation might not last for long—pipelines can usually be repaired fairly quickly—but at the very least, a painful spike in oil prices would result. The three targeted pipeline-system components need security measures reviewed at the very least. At first glance, it seems that a reasonable strategy to mitigate such attacks would add a third Abqaiq-Yanbu pipeline, parallel but not collocated for obvious reasons. However, this pipeline would extend 1200 kilometers and, estimating from other pipeline construction projects around the world, might cost one billion dollars (Pipeline & Gas Journal 2005). Clearly, other options require exploration.

4.3 The DC-Metro System: A Defender-Attacker Model

Terrorists have certainly considered the possibility of attacking the United States with nuclear, biological, or chemical (NBC) agents. In likely scenarios, terrorists contaminate a civilian population with a chemical or biological agent, or with radioactive debris from a “dirty bomb.” Subway systems in metropolitan areas seem to be attractive targets for this purpose, because their efficiency in moving large numbers of people, quickly, over long distances, would also spread a contaminant among large numbers of people, quickly, over long distances.

Consequently, authorities have already begun to install NBC sensors in the Washington, DC subway system (“DC Metro”) and in other transportation facilities around the country (Chang 2003). NBC sensors are expensive, so given limited budgets, how should these detectors be deployed? Figure 3 displays a diagram of the DC Metro, and depicts optimal locations given a supply of three sensors. “Optimal” implies that the locations minimize the worst-case time to detection (i.e., no matter where a terrorist might strike). By minimizing detection time, trains could be stopped as quickly as possible after an attack and hazardous-material response teams called in to help reduce casualties. The detection-time objective function only takes transit times and inter-platform transfer times into account, but it could certainly account for passenger volumes, if desired.

We will not provide details of this min-max defender-attacker model, but we note that related models have been studied for detecting malevolent contamination of a municipal water system (e.g., Berry et al. 2004).

Figure 4 shows the value of the optimal solution for varying numbers of detectors. This diagram leads to the key insight from this analysis, an insight that a policy-maker can easily understand: deaths versus dollars.

Before leaving this topic, we must add a caveat, lest the reader be lulled into a false sense of security. At this stage in the development of NBC detectors, especially biological detectors, noxious substances cannot be quickly and reliably identified. Such detectors may be able to identify a “suspicious” substance instantaneously, but determining if suspicion is fact may take many hours; sensitivity must be increased and false positives decreased if such technology is to prove useful. A Defense Science Board report states “...in fact, a technological breakthrough is needed.” (Defense Science Board 2004).

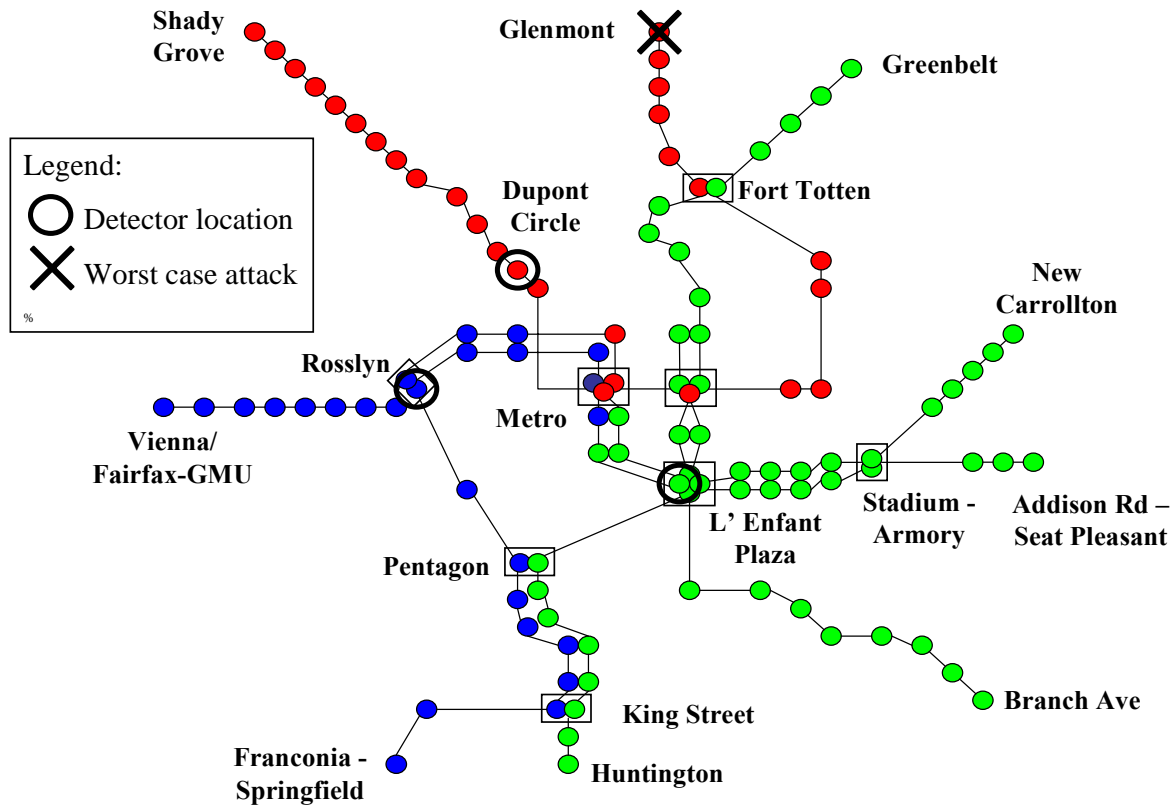


Figure 3: Locations of nuclear-biological-chemical (NBC) detectors in the Washington, D.C. Metro (subway) System to minimize maximum time to detect an attack. Using public Metro maps and schedules, we model the circulation of an NBC agent throughout the network. The solution installs detectors at Dupont Circle, L'Enfant Plaza, and Rosslyn. Observing this, an optimizing attacker would choose Glenmont to maximize the time to first detection: 31 minutes. (Avital et al. 2005).

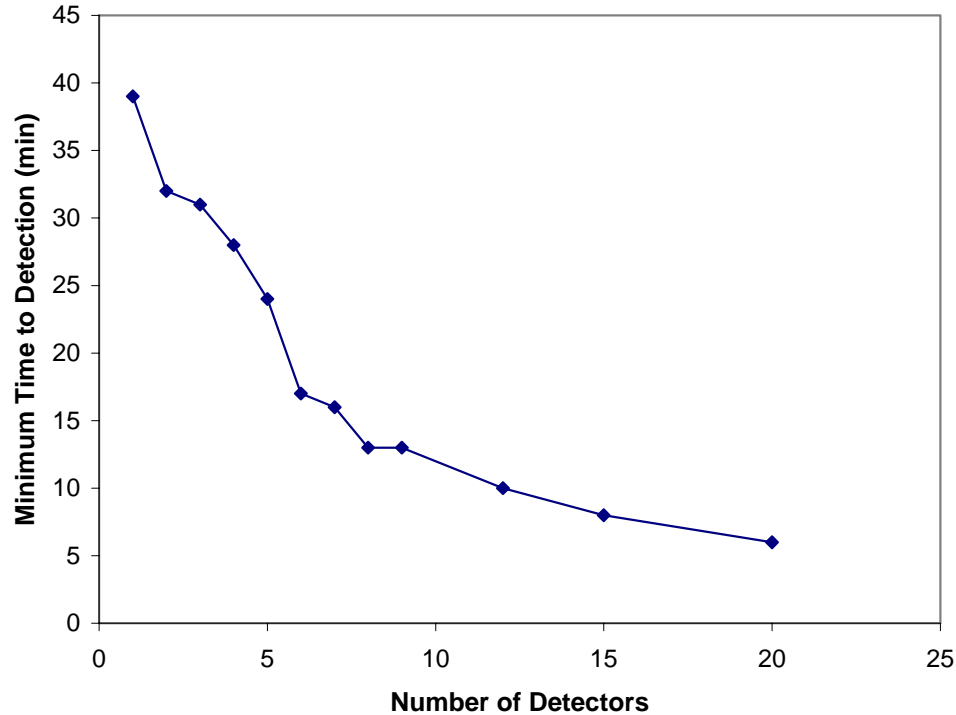


Figure 4: Minimum detection time of a nuclear-biological-chemical (NBC) attack on the DC Metro as a function of the number of detectors. This is a display for policy-makers: The horizontal axis converts to investment of millions of dollars, and the vertical axis converts to the number of people exposed to NBC contaminants, i.e., potential casualties. An analysis like this frequently reveals sharp break points (say, at 6 detectors here) that may affect policy-making (Avital et al. 2003).

4.4 Improving Airport Security: A Defender-Attacker Model

Airport security has received much attention in recent years, mostly regarding the effectiveness, or ineffectiveness, of personnel and equipment at security checkpoints (Miller 2005). But, the system aspects of airport security deserve the attention of OR analysts. Here, we investigate techniques to improve the probability of detecting a terrorist, who is trying to: infiltrate Terminal One at the Los Angeles International Airport (LAX); reach an airline gate; and hijack or sabotage an airplane. For simplicity, we consider only a single terrorist, or

“infiltrator,” who moves along the standard paths that legitimate passengers use.

Figure 5 shows a map of Terminal One, along with a skeleton of the “infiltration network” that describes the paths that an infiltrator could take from “curbside,” into the terminal, through a check-in procedure, through one or more security checkpoints, and finally out to the airline gates. (The full network contains too many arcs to depict.) We shall represent the airport’s administration: Our goal is to spend a limited “defense budget” on screening devices and procedures that increase detection probabilities on individual arcs, with the purpose of maximizing the overall detection probability. The options for changing procedures include, for instance, simply closing off certain ingress routes, or performing a physical search of, say, every third passenger, rather than every tenth. In addition to improving standard screening equipment, the red team analyzing this scenario (Landon, Takagi and Watts 2004) also includes the potential installation of advanced imaging devices now undergoing field tests (Levine 2005). As in all our models, we assume information transparency.

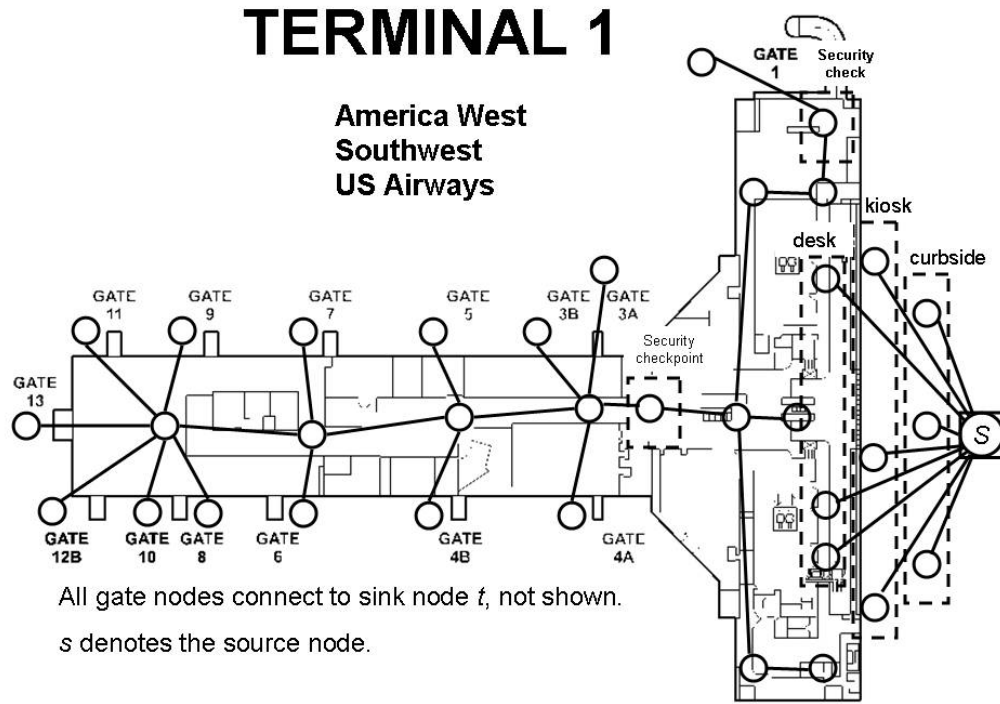


Figure 5: A limited security budget can be optimally allocated to protect Los Angeles International Airport (LAX) Terminal 1. This figure displays a map of the terminal along with a skeleton of the “infiltration network” that represents infiltration routes for terrorists (and the routes the legitimate passengers use). Arcs not shown represent movements from check-in desks or automated check-in kiosks to screening stations, through screening stations, through physical-search stations, and also artificial arcs connecting each gate node to a single sink node t . (Landon, Takagi and Watts 2004)

Probability of non-detection proves to be a more convenient concept with which to describe a defender-attacker model for this problem. For simplicity, we assume every arc k in the network possesses some nominal probability of non-detection, $q_k > 0$. This is the probability the infiltrator will not be detected if he traverses arc k . If we spend exactly c_k dollars at arc k , a new device will be installed, or a new procedure implemented, and the non-detection probability becomes $\bar{q}_k > 0$, with $\bar{q}_k < q_k$. (Notes: (i) The model extends easily to handle multiple options for reducing non-detection probability on an arc, (ii) completely closing off a route can be

handled by setting \bar{q}_k arbitrarily close to 0, and (iii) every artificial arc k connecting t in \mathcal{G} has $q_k = \bar{q}_k = 1$.) Our overall task is to expend a total budget of c' dollars so as to maximize the minimum probability of non-detection along any path the infiltrator might take. Assuming independence of detection events, this model can be formulated as follows (see the related stochastic-programming model in Pan et al. 2003):

Indices and Structural Data:

- $i \in \mathcal{N}$ nodes of the infiltration network
- $k \in \mathcal{A}$ directed arcs of the infiltration network
- $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ infiltration network

Variables:

$$x_k = \begin{cases} 1 & \text{if the defender upgrades security on arc } k \\ 0 & \text{otherwise} \end{cases}$$

$$y_k = \begin{cases} 1 & \text{if the attacker traverses arc } k \text{ when } x_k = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{y}_k = \begin{cases} 1 & \text{if the attacker traverses arc } k \text{ when } x_k = 1 \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{x}, \mathbf{y}, \bar{\mathbf{y}}$ vector forms of x_k, y_k , and \bar{y}_k , respectively

Data:

- A node-arc incidence matrix corresponding to \mathcal{G}
- \mathbf{b} node-length vector with $b_s = 1, b_t = -1$ and $b_i = 0$ for all $i \in \mathcal{N} - s - t$
- q_k nominal probability of non-detection on arc k when $x_k = 0$ ($q_k > 0$, vector form \mathbf{q})

- \bar{q}_k probability of non-detection on arc k when $x_k = 1$ ($q_k > \bar{q}_k > 0$, vector form $\bar{\mathbf{q}}$)
- d_k $\ln q_k$ (natural log of q_k) (vector form \mathbf{d} , matrix form $D = \text{diag}(\mathbf{d})$)
- \bar{d}_k $\ln \bar{q}_k$ (vector form $\bar{\mathbf{d}}$, matrix form $D = \text{diag}(\bar{\mathbf{d}})$)
- c_k cost, in dollars, to upgrade security on arc k (vector form \mathbf{c})
- c' total budget, in dollars, for upgrading security

Formulation: **See Errata, Note 2**

$$(\text{DA}_{\text{LAX}}1) \quad \min_{\mathbf{x} \in X} \max_{\mathbf{y}, \bar{\mathbf{y}}} \prod_{k \in \mathcal{A}} q_k^{(1-x_k)y_k} \bar{q}_k^{x_k \bar{y}_k} \quad (1.14)$$

$$\text{s.t.} \quad A\mathbf{y} + A\bar{\mathbf{y}} = \mathbf{b} \quad (1.15)$$

$$\mathbf{y}, \bar{\mathbf{y}} \in \{0, 1\}^{|\mathcal{A}|} \quad (1.16)$$

where $X = \{ \mathbf{x} \in \{0, 1\}^{|\mathcal{A}|} \mid \mathbf{c}\mathbf{x} \leq c' \}$.

Constraints (1.15) and (1.16) ensure that one unit of “unsplittable flow,” representing the infiltrator, moves from s to t . Constraints (1.15) are standard flow-balance constraints, just like those one could use to model a shortest-path problem in $\mathcal{G}' = (\mathcal{N}, \mathcal{A} \cup \bar{\mathcal{A}})$, which is simply \mathcal{G} with each arc duplicated.

The standard reformulation technique for this model takes a logarithm of the objective function, say the natural logarithm. This leads to the essentially equivalent model, $\text{D}_{\text{LAX}}2$, below. It is clear then that we can replace constraints (1.16) with simple non-negativity restrictions, because the constraint matrix (1.15) is totally unimodular, and for fixed \mathbf{x} , the model defines a straightforward shortest-path problem on \mathcal{G}' (if we multiply the non-positive objective function by -1 , and switch the inner maximization to a minimization). Note also that the

infiltrator's objective can only worsen by putting flow around a cycle, so no difficulty analogous to negative-length cycles in a shortest-path problem can arise.

$$\begin{aligned}
 (\text{DA}_{\text{LAX}} 2) \quad & \min_{\mathbf{x} \in X} \max_{\mathbf{y}, \bar{\mathbf{y}}} (\mathbf{1} + \mathbf{x})^T D \mathbf{y} + \mathbf{x}^T \bar{D} \bar{\mathbf{y}} && \text{See Errata, Note 3} \\
 \text{s.t.} \quad & A \mathbf{y} + \bar{A} \bar{\mathbf{y}} = \mathbf{b} \\
 & \mathbf{y}, \bar{\mathbf{y}} \geq \mathbf{0}
 \end{aligned}$$

Clearly, this model converts easily into an MILP.

Before reporting computational results, we note that modifying security equipment and procedures can both increase delays for legitimate passengers, or decrease them. For instance, increasing the percentage of people receiving physical searches on an arc will certainly increase the detection probability for an infiltrator traversing that arc, but it will also raise the average passenger's delay there. On the other hand, adding parallel metal detectors, parallel imaging devices and parallel personnel to oversee this equipment will reduce average delays. $\text{DA}_{\text{LAX}} 2$ can be modified to incorporate constraints that limit, at least approximately, the average delay for a legitimate passenger. However, for simplicity, we simply report the changes in delay that result from changes in security procedures and equipment, under the pessimistic assumption that passengers do not adjust their routes to reduce delay for themselves.

For obvious reasons, our red team can only make educated guesses about the cost of, and improved detection probabilities for, these devices. The team must also make similar guesses regarding the delay that new imaging devices will cause passengers. Therefore, the absolute statistics reported by the team cannot be taken literally. But, the relative results are believable, and the methodology can accept any system-describing parameters, which field testers and manufacturers should eventually be able to provide. We summarize the red team's computation

results below. Note that “Risk” reflects probability of detection only as a relative value, and the expenditures are probably optimistic and the delay values are probably pessimistic:

1. Baseline, scenario 1, no security improvements: Budget = \$0, “Risk” (to the infiltrator) = 10, Delay (incremental) = 0 hours, Actions = {}.
2. Scenario 2: Budget = \$100,000, Risk = 126, Average Delay = 1.5 hours, Actions = {Add two imaging devices, screen 1 in 10 at two locations, close three check-in kiosks}. (Note: Closing an automated kiosk increases the reliability of identification checks.)
3. Scenario 3: Budget = \$250,000, Risk = 249, Average Delay = 2.5 hours, Actions = {Add 15 imaging devices, screen 1 in 3 at all locations security checkpoints, close three check-in kiosks}.

4.5 Supply Chains

Supply chains, i.e., physical-distribution systems, are a key infrastructure of private-sector companies that manufacture and/or distribute goods. “Supply chains” do not appear on the list of critical infrastructure systems shown in Table 1, but they are certainly critical to our nation’s well-being.

Strategic supply-chain design has a long and successful record in the US, reducing costs and increasing service levels. Unfortunately, efficient supply chains are fragile. In fact, after scrupulously investing exactly the right amount of money in a supply chain, on exactly the right bottlenecks, the resulting product-flow patterns resemble one or more spanning trees. But, as any OR analyst knows, a spanning tree is maximally fragile: Breaking any link disconnects the network.

To address supply-chain vulnerability, we have teamed with Prof. Terry Harrison of Pennsylvania State University and Dr. Jeffrey Karrenbauer, President of INSIGHT, Inc., a company devoted to supply-chain optimization for over 25 years (INSIGHT 2004). Together, we have analyzed detailed corporate supply-chain data for many companies, including the majority of the FORTUNE 50. And, we have developed new features for INSIGHT's supply-chain optimization tools to evaluate and mitigate supply-chain vulnerability. Many key results have already been presented by Brown et al. (2003a,b) and by Brown, Carlyle, Harrison, et al. (2004), so we provide only an overview here.

The first key "result" is an observation: We still encounter considerable confusion in the private sector between random acts of nature—these have been studied by insurance actuaries for centuries—and belligerent acts of intelligent terrorists who observe defensive preparations and act to maximize damage. We strongly suggest remedying this confusion before proceeding with any analysis.

On occasion, one can reduce vulnerability substantially with simple planning and with only a modest investment in new physical infrastructure. Sometimes, just strategically relocating spare capacity—nominally used to protect against unforeseen demand surges—can provide benefit at virtually no cost. This contrasts with the high cost of adding redundant capacity, or hardening components, in other types of infrastructure such as pipelines and electric power grids.

We have learned that labor unions and competitors can be just as clever and determined as terrorists, and have similar goals: maximize damage inflicted (to market share, profit, reputation, etc.). The denial of access to west-coast ports in the US in 2002, due to a labor dispute, was no less damaging than the anthrax attacks of 2001 that closed postal and shipping services on the east coast.

We have presented our findings to numerous companies, with enthusiastic responses to even simple discoveries. American companies now have senior executives focused on “preserving corporate continuity.” These positions were originally motivated by threats to information systems, and thus back-up computer facilities and doubly backed-up data have become ubiquitous. Now, these same companies are coming to realize that they must also back up their physical operations to handle attacks on their own infrastructure (e.g., equipment, warehouses) as well as on public infrastructure they use (e.g., roads, communications networks).

4.6 Other Systems

Our work on critical infrastructure protection represents just one aspect of a research program that has also led to new military and diplomatic planning models; two have already been incorporated into comprehensive decision-support systems. One system helps plan theater ballistic-missile defense (Brown et al. 2004). The embedded defender-attacker model optimally locates anti-missile platforms (ships or ground-based units supplied with anti-missile missiles) while assuming the attacker can see some or all of our defensive preparations. The other system identifies optimal actions (e.g., embargoes of key materials, economic sanctions, military strikes) to delay a covert nuclear weapons program (Brown et al. 2005). This is an attacker-defender model where we, for a change, are the attacker. As with the missile-defense model, analysis can be carried out under different assumptions regarding the adversary’s (defender’s) ability to observe our actions. This model applies to any complex industrial project that can be delayed by a competitor.

A key insight from these military and diplomatic exercises is that deception and secrecy can make huge contributions to successful defense of our critical assets, or to successful attacks on an adversary’s critical assets. (The techniques of two-person game theory can also be useful

here; for example, see Owen 2001, pp.11-31.) Another important discovery is that the most critical tasks in an adversary's project are not obvious. For instance, decision makers with non-technical backgrounds may jump to the conclusion that the best single task to delay must lie on the nominally critical path; of course it need not.

Although this work is all relatively new, there is already an emerging body of unclassified publications including about fifty case studies, over twenty graduate theses, open-literature publications, and a number of prototypic decision-support tools. Table 3 shows some of the topics studied. We are working with certain institutions that address these threats and welcome inquiries from others. We also provide classified products to planners when the need arises.

| | |
|---------------------------------------|---------------------------------------|
| Electric grids | Weapons of mass destruction (WMDs) at |
| Road networks | Reliant Stadium, Houston, Texas |
| Strategic rail networks | Washington, DC Metro |
| Domestic water systems | Melbourne, Australia |
| Sea lanes, canals, restricted straits | Changi Naval Base, Singapore |
| Multicommodity supply chains | Manhattan |
| Petroleum distribution networks for | Norfolk, Virginia |
| US Southwest | Insurgent incursions |
| Northern California | Economic warfare |
| Defense Fuel Supply System, Japan | WMD development project |
| Theater ballistic missile attacks | |

Table 3. Case studies of these topics have evaluated roles of both attacker and defender.

5 What We Have Learned

We have discovered much through our own mathematical modeling of critical-infrastructure protection, and from applications and red-team studies. We have also learned from reading the literature, attending conferences, and speaking with colleagues, clients, and students. This section summarizes the lessons we have gleaned from all these sources.

The attacker has the advantage. This is the reverse of classical military theory and accrues from the asymmetric nature of this conflict: The defender must protect a huge, dispersed target set, while the attacker need only focus on a small set of targets chosen to maximize damage.

Some systems are naturally robust, while others are not. It turns out that our road systems are remarkably robust, fuel-distribution systems are highly fragile, and most other systems lie somewhere in between.

Hardening an infrastructure system from attack can be expensive. However, if you understand what the most damaging attacks must look like, you can better improve the system's robustness against attack for a given budget.

Critical infrastructure has been built to be "cost-effective" with little concern for coordinated, belligerent attacks. Consequently, these systems are fragile with respect to such attacks. And, even four years after September 11th, private owners of infrastructure have few economic incentives to spend large sums of money to reduce this fragility. This calls for (i) government subsidies, changes to tax codes, and regulatory reform, and/or (ii) proving the secondary economic benefit of these expenditures, if such exist (for example, spare electric transmission capacity could provide new, profitable trading opportunities).

The data are out there, and if we can get them, anybody can. "Sunshine laws" in the United States require that our governments, federal to local, conduct their affairs with transparency to the public. One result of these laws is that government agencies have produced lots of excellent web sites with lots of useful information for terrorists, based anywhere in the world. Many web sites have been redesigned in recent years to reduce access to potentially dangerous information, but we find stunning exceptions. We advise owners of public websites

associated with infrastructure to appoint an independent “red team” to analyze the website with intent to cause harm to the owners or to the users of the infrastructure.

The answers are not always obvious. The most damaging coordinated attacks, or the most effective defenses, can be non-intuitive. Key US infrastructure systems are huge, and analysis at large scale deserves rigorous, purpose-built decision-support tools to formalize the notion of a transparent, two-sided conflict. Analysis using heuristics or some other means is probably better than no analysis, but we have demonstrated how poor “reasonable” heuristics can be. (And, although not reported in this paper, we have been surprised at how poorly some of our own “clever” heuristics perform.)

Malicious, coordinated attacks can be much more damaging than random acts of nature. Our audiences usually arrive with the opposite point of view. Yes, skillful, small-scale attack can inflict more damage than a major hurricane or earthquake.

Reliability is not the answer. We must protect the most *critical* components in our infrastructure systems, rather than backing up the least reliable components. Many infrastructure owners still think that a “reliable system,” i.e., a system that fails rarely due to random events, will be a “robust system” in the face of malicious, coordinated attacks. But, common sense (for a terrorist) dictates: Destroy the most reliable components. After all, they have been made most reliable because they are most critical for system operations.

The right redundancy may be the answer. For any given level of investment, there is usually a dominant set of incremental changes to infrastructure that returns maximal immediate benefit. For some types of infrastructure, e.g., supply chains, benefit can be achieved at relatively modest cost, by adding a few alternate shipment paths, or by installing excess capacity at just the right locations, etc.

Secrecy and deception are valuable. Our military applications of attacker-defender and defender-attacker models have shown that much can be gained from secrecy and/or deception. For instance, hiding the location of a defensive asset can cause the attacker to strike a target that is essentially invulnerable. Clearly, in the world of suicide terrorists and physical infrastructure, such an outcome could be desirable.

However, the realm of deception and secrecy also lends itself to analysis with the theory of two-person zero-sum games (e.g., Owen 2001, pp.11-31). Combining that theory, and the theory of Stackelberg games as used in this paper, some rules of thumb appear for critical infrastructure protection: Hide key information about critical infrastructure, do not advertise that you are making preparations to protect particular systems, but if you do advertise—as we must in a free society—do not be too explicit.

Worst-case analysis using optimization is key to a credible assessment of infrastructure vulnerability, and to reducing that vulnerability. We cannot depend on standard reliability analyses to protect us adequately, because we cannot assume that attacks occur randomly. We face a determined, intelligent enemy who seeks to do us maximal harm.

Acknowledgements

Salmeron and Wood thank the US Department of Homeland Security, and the US Department of Energy for research support. Brown and Wood are grateful for sustaining research support from the Office of Naval Research (Division of Mathematical Sciences), the Air Force Office of Scientific Research (Optimization and Discrete Mathematics Program), and the Joint Warfare Analysis Center. Brown, Carlyle and Wood also thank the National Security Agency for support. At various times, we have also worked with and received support from

every US uniformed military service, and we thank every one.

In late 2001, we approached INSIGHT, Inc., to help us discover what private companies could do to fortify their operations against hostile threats. (Note: Brown and Wood have worked on private-sector business optimization problems for decades with INSIGHT.) INSIGHT has granted unfettered use of its supply-chain design software, devoted extensive development effort, provided data from a host of private-sector clients (scrubbed of proprietary identification and confidential data), and arranged direct access to its clients. We are grateful for INSIGHT's assistance.

References

- Andrews, C., Cason, K., Godfrey, A. and Revor, M., 2003, "Optimal Defense of Saudi Arabia's Pipelines against Terrorist Attack," Red Team Report, OA4202 Network Flows and Graphs, Naval Postgraduate School, Monterey, California, November.
- Avery, W., Brown, G.G., Rosenkranz, J., and Wood, R.K., 1992, "Optimization of Purchase, Storage and Transmission Contracts for Natural Gas Utilities," *Operations Research*, **40**, pp. 446-462.
- Avital, I., Cormican, K. and Wood, K., 2005, "Where to Install Contamination Detectors in a Subway System," in preparation.
- Benders, J.F., 1962, "Partitioning Procedures for Solving Mixed Integer Variables Programming Problems," *Numerische Mathematik* **4**, 238-252.
- Berry, J., Fleischer, L.K., Hart, W.E. and Phillips, C.A., 2004, "Sensor Placement in Municipal Water Networks," World Water and Environmental Resources Congress, Philadelphia, June 23-26, SAND2003-0918C.
- Bremmer, I., 2004, "The Saudi Paradox," *World Policy Journal*, Volume 21, No. 3, Fall, pp. 23-30.
- Brown, E., Hakola, M., Raffetto, M. and Yanik, T., 2003, "Effects of Terrorist Attacks at US Airports," Red Team Report, OA4202 Network Flows and Graphs, Naval Postgraduate School, Monterey, California, November.

Brown, G., M. Carlyle, D. Diehl, J. Kline, and K. Wood, 2004, "How to Optimize Theater Ballistic Missile Defense," *Operations Research*, to appear.

Brown, G., Carlyle, M., Harney, R., Skroch, E. and Wood, K., 2005, "Interdicting a Nuclear Weapons Project," (in review).

Brown, G., Carlyle, M., Harrison, T., Salmerón, J. and Wood, K., 2003a, "How to Attack a Linear Program," plenary address, Military Operations Research Society, Quantico, VA, 10 June.

Brown, G., Carlyle, M., Harrison, T., Salmerón, J. and Wood, K., 2003b, "Tutorial: How to Build a Robust Supply Chain or Harden the One You Have," INFORMS Annual Meeting, Atlanta, GA, 19 October.

Brown, G., Carlyle, M., Harrison, T., Salmerón, J. and Wood, K., 2004, "Designing Robust Supply Chains and Hardening the Ones You Have," INFORMS Conference on OR/MS Practice, Cambridge, MA, 26-27 April.

Brown, G.G., Dell, R.F., and Wood, R.K., 1997, "Optimization and Persistence," *Interfaces*, **27**, pp. 15-37.

Chaney, E. and Berner, R. 2004, "Global: Oil Price Update: Still Higher and More Uncertain," *Global Economic Forum*, Morgan Stanley, 4 May, <http://www.morganstanley.com/GEFdata/digests/20040507-fri.html>, accessed 7 May 2005.

Chang, K., 2003, "Ideal Sensors for Terror Attack Don't Exist Yet," **New York Times**, 1 April, p. F1.

Cormican, K. J., Morton, D. P. and Wood, R. K., 1998, "Stochastic Network Interdiction," *Operations Research*, **46**, 184-197.

Defense Science Board, 2004, "Defense Science Board 2003 Summer Study on DoD Roles and Missions in Homeland Security," Volume II-A: Supporting Reports, Office of the UnderSecretary of Defense for Acquisition, Technology, and Logistics, Washington, DC, May.

Department of the Army, 2000a, *Army Field Manual FM 3-01.11*, Appendix A – ADA Employment Principles, Guidelines, and Priorities, <http://www.globalsecurity.org/military/library/policy/army/fm/>, accessed 18 April 2005.

Department of the Army, 2000b, *Army Field Manual FM 44-100*, Chapter 4. Fundamentals of Army Air and Missile Defense Operations, <http://www.globalsecurity.org/military/library/policy/army/fm/>, accessed 18 April 2005.

Department of Homeland Security 2002, National Strategy for Homeland Security, <http://www.whitehouse.gov/homeland/book/>, accessed 18 April 2005.

Department of Justice, 2004, "Al Qaeda Training Manual," <http://www.usdoj.gov/ag/trainingmanual.htm> , accessed 18 April 2005.

Garcia, M.L., 2001, *The Design and Evaluation of Physical Protection Systems*, Butterworth-Heinemann.

Golden, B., 1978, "A Problem in Network Interdiction," *Naval Research Logistics Quarterly*, **25**, 711-713.

Insight, 2004, Strategic Analysis of Integrated Logistics Systems (SAILS), Manassas, VA, <http://www.insight-mss.com> , accessed 18 April 2005.

Israeli, E. and Wood, K., 2002, "Shortest-Path Network Interdiction ," *Networks*, Vol. 40, pp. 97-111.

Landon, C., Takagi, K. and Watts, K., 2004, "Protection of Restricted Areas from Infiltration," Red Team Report, OA4202 Network Flows and Graphs, Naval Postgraduate School, Monterey, California, November.

Levine, S., 2005, "Software and Smart Card," US News & World Report, 14 March, <http://www.usnews.com/usnews/news/articles/050314/14airport.htm>, accessed on 9 May 2005.

Miller, L., 2005, "Airport Security Hasn't Improved," Associated Press, 16 April.

Moore, J. T. and Bard, J. F., 1990, "The Mixed Integer Linear Bilevel Programming Problem," *Operations Research*, 38, pp. 911-921.

NORTHCOM 2004, <http://www.northcom.mil/>, accessed 18 April 2005.

Pan, F., Charlton, W.S. and Morton, D.P. , 2003, "A Stochastic Program for Interdicting Smuggled Nuclear Material," in *Network Interdiction and Stochastic Integer Programming*, D.L. Woodruff (ed.), Kluwer Academic Publishers, pp. 1-20.

Pipeline & Gas Journal, 2005, "Pipeline Construction Scorecard," January, pp 82-85.

Owen, G., 2001, *Game Theory*, Third Edition, Academic Press, San Diego, California.

Roberts, N.H., Vesely, W.E., Haasl, D.F. and Goldberg, F.F., 1981, *Fault Tree Handbook*, NUREG-0492, U.S. Nuclear Regulatory Commission, Washington, DC.

Salmerón, J., Wood, R.K. and Baldick, R., 2004a, "Optimizing an Electric Grid Under Asymmetric Threat (II)," Technical Report NPS-OR-04-001, Naval Postgraduate School. <http://library.nps.navy.mil/uhtbin/hyperion/NPS-OR-04-001.pdf>

Salmerón, J., Wood, K. and Baldick, R., 2004b, "Analysis of Electric Grid Security Under

Terrorist Threat,” *IEEE Transactions on Power Systems*, Vol. 19-2, pp. 905-912.

Salmerón, J., Wood, R.K. and Baldick, R., 2005, “VEGA,”
<http://www.nps.navy.mil/orfacpag/resumePages/projects/VEGA.htm>, accessed 18 April 2005.

Sparshott, J., 2004, “Saudi’s Security for Oil Extensive,” **The Washington Times**, 22 June.

US-Canada Power System Outage Task Force, 2003, “Causes of the August 14th Blackout in the United States and Canada,” interim report. <http://www.nerc.com/~filez/blackout.html>, accessed 18 April 2005.

von Stackelberg, H., 1952, *The Theory of the Market Economy*, (trans. from German), William Hodge & Co., London.

Wallace, B., 1998, “3-Year Term for Army Vet Who Fired On Power Plant,” **San Francisco Chronicle**, 24 April, p. A21.

Washington Post 2002, “Sniper Shootings,”
<http://www.washingtonpost.com/wp-srv/metro/daily/oct02/snipershootings.htm>, accessed 18 April 2005

Wood, R.K., 1993, “Deterministic Network Interdiction,” *Mathematical and Computer Modelling*, 17, pp. 1-18.

Wood, A.J. and Wollenberg, B.F., 1996, *Power Generation, Operation and Control*. Second Edition. John Wiley and Sons, New York.

Errata

G. Brown, M. Carlyle, J. Salmeron, and K. Wood, 2005, “Analyzing the Vulnerability of Critical Infrastructure to Attack , and Planning Defenses,” in *Tutorials in Operations Research* , INFORMS, pp. 102-123.

(The page numbers and equation numbers are different in the published version of the paper, but the same errors appear there as in this draft.)

Note 1, Page 11 of the draft:

In (AD1-MILP),

$$\text{s.t. } A^T \boldsymbol{\theta} + F^T \boldsymbol{\beta} - F^T P \mathbf{x} \leq \mathbf{c}$$

should be

$$\text{s.t. } A^T \boldsymbol{\theta} + F^T \boldsymbol{\beta} - F^T P \mathbf{x} \leq \mathbf{c}^T .$$

Note 2, Page 33 of the draft:

(DA_{LAX}1) on page 33 should be

$$\begin{aligned} (\text{DA}_{\text{LAX}} 1) \quad & \min_{\mathbf{x} \in X} \max_{\mathbf{y}} \prod_{k \in \mathcal{A}} (q_k^{(1-x_k)} \bar{q}_k^{x_k})^{y_k} \\ & \text{s.t. } A\mathbf{y} = \mathbf{b} \end{aligned} \tag{1.15}$$

$$\mathbf{y} \in \{0,1\}^{|\mathcal{A}|} \tag{1.16}$$

$$\text{where } X = \left\{ \mathbf{x} \in \{0,1\}^{|\mathcal{A}|} \mid \mathbf{c}\mathbf{x} \leq c' \right\} .$$

Note 3, Page 33 of the draft:

(DA_{LAX}2), on page 34, should be

$$\begin{aligned} (\text{DA}_{\text{LAX}} 2) \quad & \min_{\mathbf{x} \in X} \max_{\mathbf{y} \geq 0} \left(\mathbf{d} + \mathbf{x}^T (\bar{D} - D) \right) \mathbf{y} \\ & \text{s.t.} \quad A\mathbf{y} = \mathbf{b}. \end{aligned}$$